Asian Journal of Applied Science and Technology (AJAST)
Volume 8, Issue 2, Pages 81-90, April-June 2024

Securing Medicinal Plant Supply Chain through ML-Enhanced Image Processing

Jemi J.A. Benisha[1*], Jenika S.[2], Manju K.M.[3] & Mrs. Jino Shiny V.[4]

[1,2,3]UG Student, [4]Assistant Professor, [1-4]Department of Electronics and Communication Engineering, Stella Mary's College of Engineering, Kanyakumari, Tamil Nadu, India. Corresponding Author (Jemi J.A. Benisha) Email: jemibenisha100@gmail.com[*]

DOI: https://doi.org/10.38177/ajast.2024.8209

Article Received: 12 March 2024 *Article Accepted: 15 May 2024* *Article Published: 19 May 2024*

ABSTRACT

The medicinal plant industry encounters numerous challenges, including misidentification and counterfeit goods. These issues have the potential to compromise the quality standard and, consequently, compromise the safety of medical products. By employing CNN, a subset of Python-based ML algorithms, this article presents a novel approach to the problem of pharmaceutical image recognition in the medicinal plant supply chain. The developed system will be predicated on a large database of images of medicinal plants, which will be utilized to instruct CNN models that have undergone rigorous training to accurately identify and authenticate plant materials throughout the entire supply chain. By utilizing an ML system for image processing, this system provides the most reliable assurance that medicinal plants are genuine and of high quality. The experimental results provide evidence that the proposed solution is valuable, as it holds potential for enhancing supply chain security, eliminating counterfeit products, and facilitating the identification of medicinal plants. The Python implementation of the system's solution facilitates user customization and integration. That would enable stakeholders to effectively address their requirements, consequently facilitating prompt decision-making. The purpose of this project is to present a specific ML and CNN technique that addresses the issue of counterfeit goods and contributes to the chain security of medical plants.

Keywords: Medicinal plant; Supply chain; Machine learning; Python; Image Processing.

1. Introduction

The natural medicine plant industry has a twofold role: firstly, it offers practical alternatives for various challenges; and secondly, it contributes to the generation of revenue for the pharmaceutical industry [1]. The main concern lies in the presence of counterfeit drugs and impurities, which not only undermine the integrity and structure of medicinal plants, but also compromise their security levels. These dangers undermine the health and diminish confidence in herbal medicines. Fraudulent herbal remedies diminish efficacy not only by compromising the quality of the product but also by deceptively incorporating dangerous components [2]. The deliberate or accidental modification of the strength and safety of plant drugs by adulteration determines the specific therapeutic effects of medicinal plants. Furthermore, the lack of proper supervision in identifying plant material may lead to the inclusion of incorrect or dangerous species, posing significant health risks to consumers [3]. It is imperative to have strong methods in place to verify and track the sources of medicinal plants at every step of the distribution process in order to tackle these difficulties.

Traditional methods of plant authentication, such as chemical analysis and visual inspection, often require a lot of time and effort, are costly, and are susceptible to mistakes made by humans [4]. As a result, continuous progress in technology is made to achieve two goals: guaranteeing the ability to track the supply of herbs and improving overall operational efficiency. Artificial intelligence technology has advanced rapidly in the past two years, allowing it to be used as a tool for verifying the integrity of supply chains. Convolutional neural networks (CNNs) are a powerful deep learning algorithm commonly used for plant species identification [5]. Their remarkable accuracy in image recognition has been widely recognized in the field. A CNN that has undergone training will have the ability to precisely classify different plant species owing to the extensive number of plant images that need to be processed

for this purpose. Therefore, they are the most suitable tool for verifying the signatures of naturally occurring plants with medicinal properties [6]. The article presents a novel approach to protecting a series of medicinal plants by combining Convolutional Neural Networks (CNN) with machine learning and Python for the purpose of image recognition. The proposed technology combines machine learning and advanced image processing to accurately identify and authenticate medicinal plants at different points along the supply chain. Unlike other methodologies that depend on human labor and subjective observation, the proposed approach is a dependable and automated technological system for plant identification. Furthermore, by harnessing these advantages, the Python programming language can easily be adapted to accommodate various environments and specific needs. The proposed model aims to combine machine learning algorithms and image recognition capabilities to accurately and comprehensively identify counterfeit products [7]. Furthermore, the main objective of the system is to verify the safety and purity of medications obtained from herbal sources, in addition to the previously mentioned factors. The following sections will explore the technical complexities of the proposed system: The essential elements of an image processing system based on machine learning are the creation and execution. Moreover, this text will clarify the process of experimentation and the resulting outcomes, illustrating that the suggested solution has the capacity to be impeccable and reliable in protecting the supply chain of medicinal plants.

2. Literature Survey

The convergence of machine learning, image processing, and blockchain technology has become a promising area of study in the field of medicinal plant authentication and traceability in recent years. This literature review examines prominent research that showcases cutting-edge approaches and technological progressions designed to tackle the obstacles encountered in the medicinal plant supply chain [8].

Sharma et al. introduced an innovative method for identifying medicinal plants by employing Convolutional Neural Networks (CNNs) based on deep learning [9]. The study guarantees accurate species identification by training Convolutional Neural Network (CNN) models using a comprehensive database of plant images. This methodology represents a substantial advancement in the precision and effectiveness of plant recognition systems, offering improved dependability in the authentication of medicinal plants.

Yadav et al. investigate the utilization of machine learning methodologies, such as support vector machines and random forests, for validating the genuineness of medicinal plants [10]. The study showcases the efficacy of machine learning algorithms in accurately identifying plant species by analyzing leaf images, thereby enhancing trust in the integrity of the supply chain. This approach not only improves the ability to track and trace products, but also makes it easier to identify products that have been tampered with, which helps ensure the quality of herbal medicine production.

Lei et al. propose the incorporation of blockchain technology to create a supply chain for medicinal plants that is both transparent and traceable [11]. By utilizing track-and-trace systems enabled by blockchain technology, each step of the plant's life cycle, starting from cultivation and ending with distribution, can be verified and made unchangeable. This guarantees responsibility and eradicates the possibility of tampering or deceitful actions, promoting confidence among stakeholders and consumers alike.

Gupta et al. present a novel automated plant recognition system that utilizes machine learning algorithms and image processing techniques [12]. The system effectively identifies medicinal plant species from leaf images by extracting features and utilizing classification methodologies. This technological advancement simplifies the process of verifying the authenticity of medicinal plants in the supply chain, providing a scalable solution for tracking and confirming their origin.

Wani et al. present a technique for assessing the quality of medicinal plants by employing digital image processing and machine learning [13]. The system utilizes machine learning algorithms to analyze morphological characteristics, enabling the detection of adulteration and ensuring the authenticity of herbal medicinal products. This strategy improves the security of the supply chain and protects public health by reducing the dangers posed by counterfeit or low-quality products.

The studies mentioned above highlight a clear shift towards the use of sophisticated technologies such as machine learning, image processing, and blockchain to tackle authentication and traceability issues in the medicinal plant supply chain [14]. Researchers seek to improve supply chain security, strengthen consumer trust, and maintain the integrity of herbal medicine production by adopting these innovative methodologies. In the future, further research and collaboration in this field have the potential to greatly transform the way medicinal plants are authenticated and traced.

3. Existing System

Prior to the development of advanced machine learning and image processing techniques, medicinal plant products were typically certified and tracked in the supply chain using methods that relied on visual inspection, chemical analysis, and record-keeping [15]. This is due to the fact that visual evaluation relies on human judgment, which introduces subjectivity into the authentication process and leads to inconsistent results when it comes to accurately identifying medicinal plants.

A. Consumptive of time: The conventional method of authenticity, which involves a chemical procedure analysis, necessitates a considerable amount of time (a precious resource) for testing and investigation in a laboratory.

B. Scalability Issues: There needs to be a more scalable solution than manual authentication for large-scale production and distribution of authentic products.

C. Ineffective Traceability: Medicinal plant origin and movement detection is a problem with both modern traceability and systems that rely on traditional documentation. Authenticator plant and tracking plant practices still rely on traditional methods, even though these approaches are encountering these problems. However, due to their limitations, new ways were sought after to incorporate innovative technology into the supply chain safety framework while evading the blockchain-based traceability model [16]. While they did their best to address the various authentication and traceability methods, they were unable to do away with the complexity and vulnerabilities inherent in today's supply chains. To strengthen logistics chain security and guarantee the authenticity of medicinal plant products, there was an immediate need for new approaches that could use technological advancements like process images and machine learning.

Asian Journal of Applied Science and Technology (AJAST)
Volume 8, Issue 2, Pages 81-90, April-June 2024

4. Proposed Method

The proposed methodology utilizes a range of machine learning techniques, with a particular focus on Python and the application of Convolutional Neural Networks (CNNs) for image recognition. The objective is to verify the authenticity of the medicinal plant supply chain. Several crucial steps in the process include forming and processing the data, extracting features from the data, training machine learning models, and conducting testing.

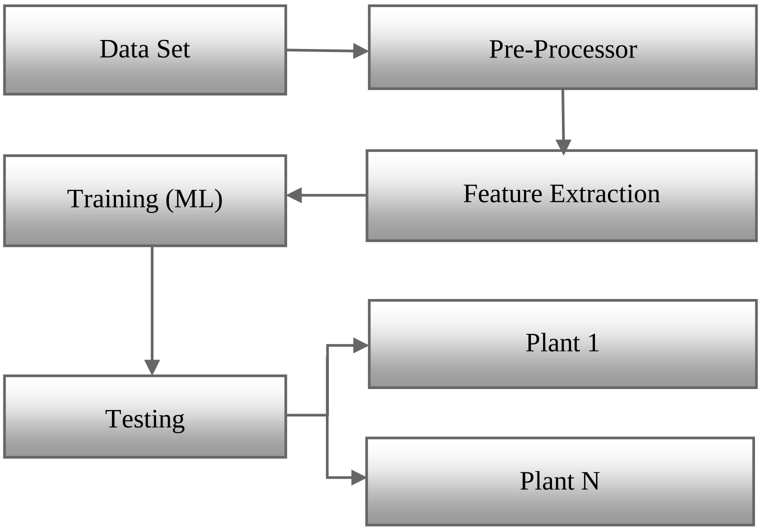

Figure 1. Machine Learning Workflow

A. Dataset Preprocessing

Generate an extensive collection of images containing numerous examples of medicinal herbs. An effective method for preparing data for a reliable predictive model involves capturing images from various backgrounds, lighting conditions, and angles. Ensure that the plant images in the dataset are accompanied by precise labels that accurately classify their identification.

B. Pre-processing

Uniformly adjust the size of all the images to ensure compatibility with CNN's image processing. This technique provides a consistent output during the model training phase, and the algorithm effectively reduces the time complexity in computing. Normalize the peak values by mapping them to a standard scale, such as [0, 1], to establish a stable foundation for training the model and enhance its efficiency. To enhance the model's generality and increase the size of the dataset, augment the sample by employing techniques such as flipping, rotating, and zooming.

C. Feature Extraction

An Embedded module's function would involve serving as a Convolutional Neural Network (CNN) trained architecture instructor in either VGG16 or ResNet50. The pre-trained model no longer includes the fully connected layers, leaving only the convolutional layers. By employing Convolutional Neural Networks (CNN), input the preprocessed images into the network's layers to extract the salient features of each image. The subsequent action involves utilizing the empirical characteristics as inputs.

D. Training Machine Learning Models

To classify medicinal plant species, it is essential to create a comprehensive framework that encompasses all the processed characteristics. To determine the probability of a class, it is necessary to incorporate one or more fully connected layers, which are subsequently followed by a softmax activation function. This function aids in obtaining the probabilities for each class. Partition the dataset into subsets that will be utilized for the purposes of validation, testing, and training. Improve the optimization of the validation set by adjusting the hyperparameters and monitoring the model's performance. Utilize the training set to facilitate the self-training of the model, and employ the testing set to evaluate the model's capacity for generalization. Enhance the classifier loss by refining the machine learning algorithm through the utilization of gradient descent and stochastic gradient descent optimization techniques.

E. Evaluating

Evaluate the performance of the trained model on the test set using metrics such as precision, recall, accuracy, and F1-score. By employing these metrics, one can gain a comprehensive comprehension of the model's functionality, enabling the creation of accurate categorizations for medicinal plants. The proposed system utilizes Convolutional Neural Network (CNN) image recognition techniques in conjunction with Python-based machine learning. Consequently, it relies on a system that is reliable and precise.

5. Methodology

The process begins with extensive data collection, where experts in the field must search for a diverse array of specialized photographs of medicinal plants that are exclusively sourced from reliable sources. Therefore, the expert in this domain has the ability to collect photographs of numerous plants that encompass various species and environmental circumstances. Identifying the labeled image for the supervised learning process is a crucial step that is linked to each species. During the pre-processing operations, the dataset undergoes resizing, normalization, and augmentation to ensure consistent and robust deployment.

Subsequently, a CNN model is selected that achieves computational efficiency while maintaining a suitable level of complexity, thus achieving a balanced approach. By initializing a pre-trained model, transfer learning can be optionally utilized to improve the model's capability of identifying medicinal plants in photos by enhancing the specificity of its features.

Secondly, when the training dataset is stratified to include the training, validation, and test sets, the model utilizes it. Performance is assessed using diverse metrics to identify patterns and promote discussed progress in accordance with the specified strategies. CNNs, or convolutional neural networks, are a commonly used AI technique for image classification tasks. The system's architecture is designed to mimic the layers of the human visual system that extract features from images in its image processing systems. Ultimately, feature maps are subjected to various pooling layers in order to reduce the amount of data to its maximum extent. The output layer generates the final predictions, while the fully connected layers handle additional feature processing and classification. The applications of Convolutional Neural Networks (CNNs), such as object detection, facial recognition, and medical

image analysis, primarily rely on the hierarchical arrangement of these structures. These structures learn intricate patterns by combining multiple layers.

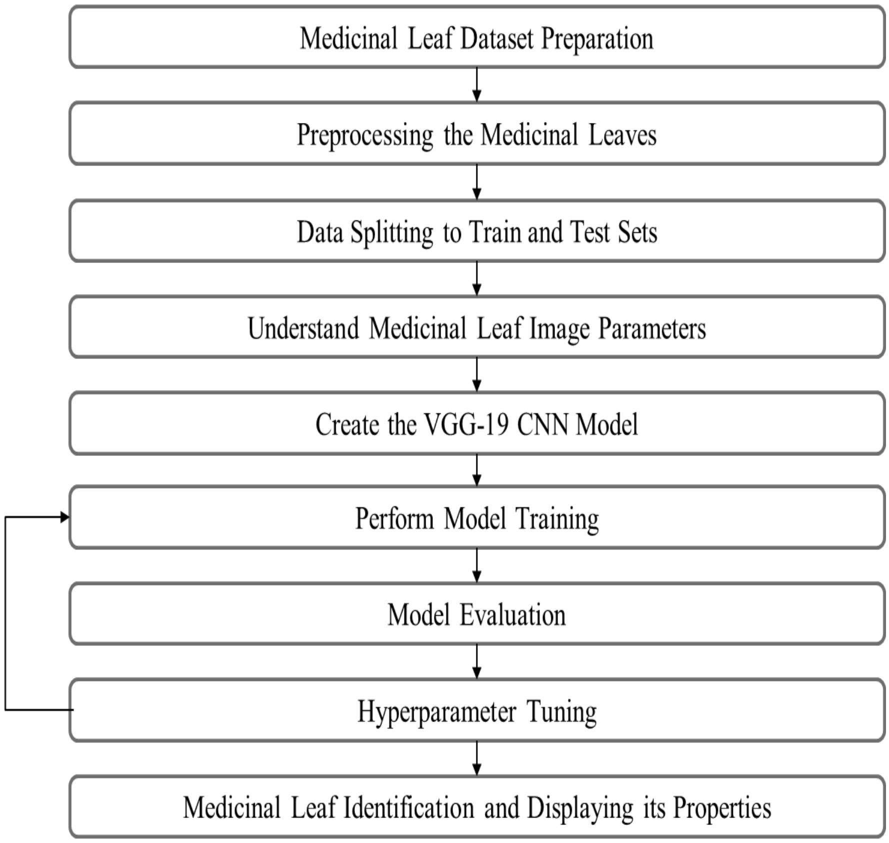

Figure 2. Data Flow Diagram

The last step involves evaluating the model by testing it on a separate dataset to determine its accuracy and ability to generalize. In addition to the analysis of the confusion matrix, we will also provide information on the specific metrics that can be utilized to address any weaknesses. The primary objectives of the fine-tuning processes are to determine the optimal values for the hyperparameters used in training kernels and to generate additional ideas for improving the performance of the models and their readiness for deployment. Ultimately, the completed model will be implemented in various real-world situations, either operating autonomously or seamlessly integrated with existing supply chain management systems. The presence of automation and monitoring systems in user interfaces and processes ensures their functionality and long-term sustainability. Consequently, stakeholders will unquestionably validate traditional medicine. We have utilized a Convolutional Neural Network (CNN) model that follows the structure of VGG19. More specifically, we have added custom classification layers to a pre-trained VGG19 model that was originally trained on the ImageNet dataset. The training data, which is inputted into the layers, helps prevent overfitting by detecting patterns in later stages.

6. Result and Discussion

The results are examined to assess the precision of the model. Initially, we begin by outlining the details of the dataset, followed by the hardware specifications. Subsequently, we introduce the models employed to evaluate and compare the outcomes achieved across different models.

Asian Journal of Applied Science and Technology (AJAST)
Volume 8, Issue 2, Pages 81-90, April-June 2024

The VGG19 architecture was chosen as the foundational model due to its proven effectiveness in image classification tasks. In order to leverage the insights acquired from a vast dataset, we utilize the weights that have been pre-trained on the ImageNet dataset.

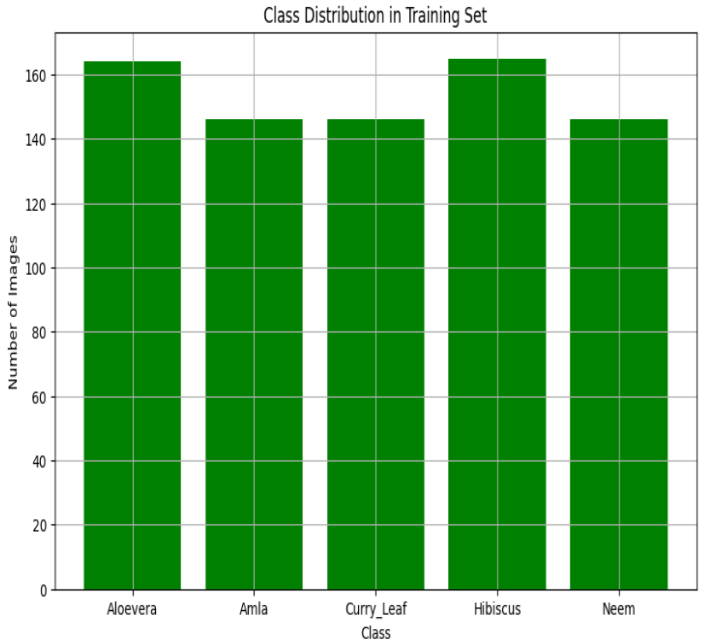

Figure 3. Class Distribution in Training Set, Reflects the number of images available for each medicinal plant images

The 'Medicinal leaf dataset' comprises 1822 images of Indian medicinal plants, classified into 30 distinct species, with each species containing 60-100 unique images. The class separation on the source data set is determined by the quantity of images corresponding to each type of medicinal plant.

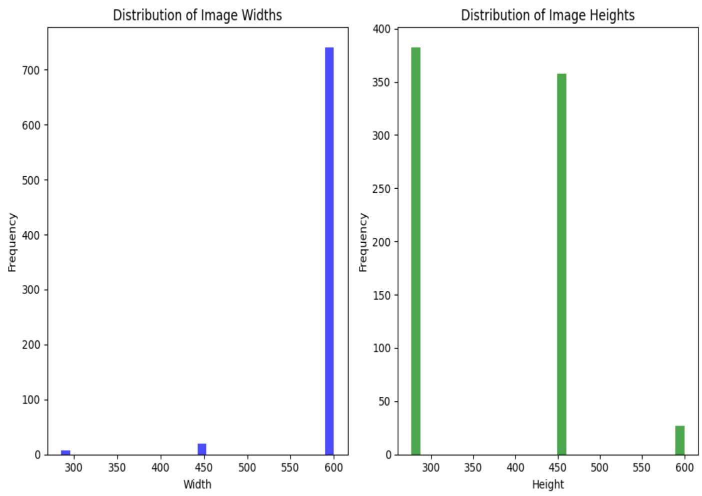

Figure 4. Distribution of image Widths and Heights

The range of the dimensions possessed by images scattered around inside the dataset is what is called the image height and width distribution. Different types of preprocessing such as resizing and standardization are also managed by knowing the images' varied scales via the analysis of this distribution.

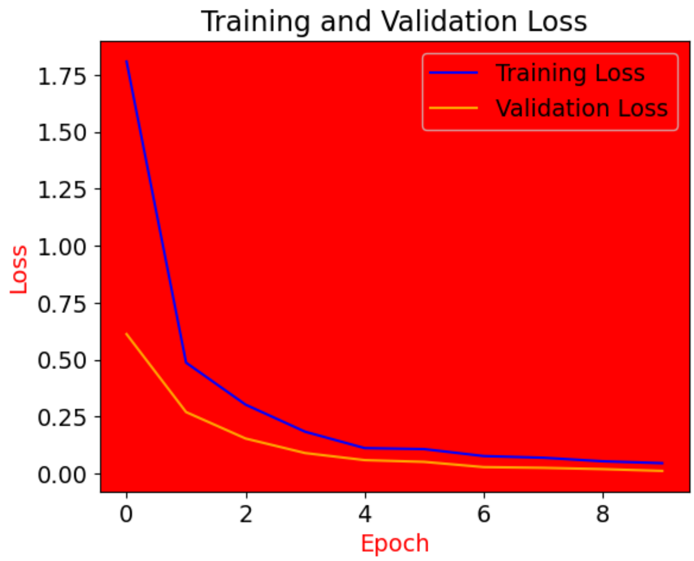

Figure 5. Training and Validation Loss

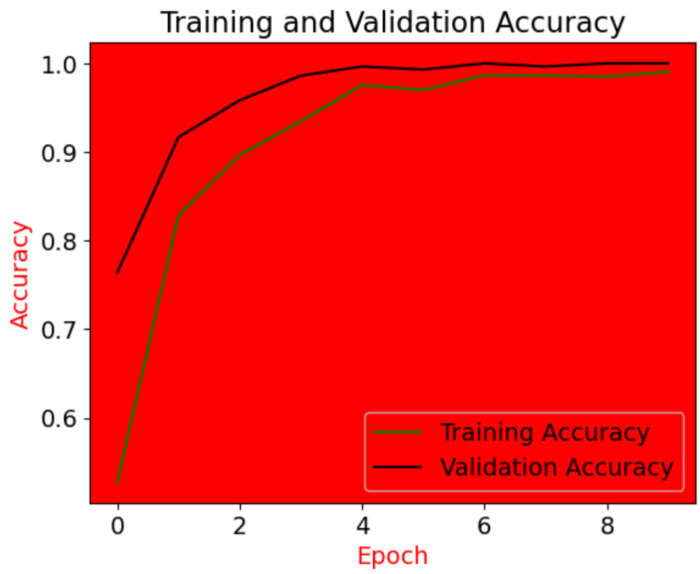

Figure 6. Training and Validation Accuracy

Now let's examine the figure where the line pattern represents the validation accuracy and the line with dots represents the validation loss. The graphs exhibit a cyclic pattern of fluctuation, wherein accuracy and loss are observed to alternate between reaching high and low values. This demonstrates the performance of the model at various stages of training. The rules can be succinctly summarized as follows: when the value curve decreases, the validation accuracy increases, and conversely.

7. Conclusion

The convergence of plant identification benchmarks utilizing the VGG19 CNN model, characterized by straightforward yet comprehensive structures, has yielded promising outcomes. The study's model analysis illustrates the efficacy of training across different classes, contributing to a refined model training process. Additionally, analyzing the distribution of image dimensions underscores the variability in image sizes, a crucial

aspect in understanding gender gap dynamics in specific fields. Validation accuracy and loss graphs affirm the model's accuracy and low loss, despite minor oscillations during training epochs, reflecting the model's learning dynamics and optimization stage. Achieving 100% validation accuracy underscores the model's proficiency in identifying plant images within the validation set. Consequently, the VGG19-CNN combination demonstrates efficacy in plant classification, with potential applications in species identification and agricultural monitoring. Continuous refinement and adaptation to diverse plant datasets promise further advancements, crucial in addressing the high demand for plant-derived drugs while mitigating environmental impacts. Machine learning and image processing technologies offer a transformative approach, enhancing therapeutic plant utilization and fostering precision agriculture, ultimately improving global well-being.

Declarations

Source of Funding

This study did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing Interests Statement

The authors declare no competing financial, professional, or personal interests.

Consent for publication

The authors declare that they consented to the publication of this study.

Authors' contributions

All the authors took part in literature review, analysis and manuscript writing equally.

References

[1] Najmi, Asim, Sadique A. Javed, Mohammed Al Bratty & Hassan A. Alhazmi (2022). Modern approaches in the discovery and development of plant-based natural products and their analogues as potential therapeutic agents. Molecules, 27(2): 349.

[2] Obahiagbon, Ebiuwa Gladys & Matthew Chidozie Ogwu (2024). Consumer Perception and Demand for Sustainable Herbal Medicine Products and Market. In Herbal Medicine Phytochemistry: Applications and Trends, Pages 1–34, Cham: Springer International Publishing.

[3] Ahen, Frederick, Kwame O. Buabeng & Outi MH Salo-Ahen (2023). Market violence through destructive entrepreneurship: Assessing institutional responses to the proliferation of counterfeit traditional and alternative medicines in Ghana. Heliyon, 9(3).

[4] Pugazhendiran, P., Suresh Kumar, K., Ananth Kumar, T., & Sundaresan, S. (2022). An Advanced Revealing and Classification System for Plant Illnesses Using Unsupervised Bayesian-based SVM Classifier and Modified HOG-ROI Algorithm. In Contemporary Issues in Communication, Cloud and Big Data Analytics: Proceedings of CCB 2020, Pages 259–269, Springer Singapore.

[5] Deshmukh, M. (2024). Deep Learning for the Classification and Recognition of Medicinal Plant Species. Indian Journal of Science and Technology, 17(11): 1070–1077.

[6] Moffat, John G., Fabien Vincent, Jonathan A. Lee, Jörg Eder & Marco Prunotto (2017). Opportunities and challenges in phenotypic drug discovery: an industry perspective. Nature reviews Drug discovery, 16(8): 531–543.

[7] Jahanbakhshi, Ahmad, Yousef Abbaspour-Gilandeh, Kobra Heidarbeigi & Mohammad Momeny (2021). A novel method based on machine vision system and deep learning to detect fraud in turmeric powder. Computers in Biology and Medicine, 136(2021): 104728.

[8] Sundaresan, S., Suresh, K., Kishore, V., & Jayakumar, A. (2021). Insight Into Various Algorithms For Medical Image Analyzes Using Convolutional Neural Networks (Deep Learning).In Handbook of Deep Learning in Biomedical Engineering and Health Informatics, Pages 137–164, Apple Academic Press.

[9] Sharma, Sanyam, Subh Naman, Jayesh Dwivedi & Ashish Baldi (n.d.). Artificial Intelligence-Based Smart Identification System Using Herbal Images: Decision Making Using Various Machine Learning Models. In Applications of Optimization and Machine Learning in Image Processing and IoT, Pages 123–155, Chapman and Hall/CRC.

[10] Yadav, Anupam (2021). Predicting covid-19 using random forest machine learning algorithm. In 2021 12th international conference on computing communication and networking technologies (ICCCNT), Pages 1–6, IEEE.

[11] Lei, Moyixi, Longqin Xu, Tonglai Liu, Shuangyin Liu & Chuanheng Sun (2022). Integration of privacy protection and blockchain-based food safety traceability: Potential and challenges. Foods, 11(5): 2262.

[12] Gupta, Deepak, Prerna Sharma, Krishna Choudhary, Kshitij Gupta, Rahul Chawla, Ashish Khanna & Victor Hugo C. de Albuquerque (2021). Artificial plant optimization algorithm to detect infected leaves using machine learning. Expert Systems, 38(6): e12501.

[13] Wani, Javaid Ahmad, Sparsh Sharma, Malik Muzamil, Suhaib Ahmed, Surbhi Sharma & Saurabh Singh (2022). Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: Methodologies, applications, and challenges. Archives of Computational Methods in Engineering, 29(1): 641–677.

[14] Ben Ayed, Rayda, Mohsen Hanana, Sezai Ercisli, Rohini Karunakaran, Ahmed Rebai & Fabienne Moreau (2022). Integration of innovative technologies in the agri-food sector: The fundamentals and practical case of DNA-based traceability of olives from fruit to oil. Plants, 11(9): 1230.

[15] Kshetri, Nir (2021). Blockchain and sustainable supply chain management in developing countries. International Journal of Information Management, 60: 102376.

[16] Centobelli, Piera, Roberto Cerchione, Pasquale Del Vecchio, Eugenio Oropallo & Giustina Secundo (2022). Blockchain technology for bridging trust, traceability and transparency in circular supply chain. Information & Management, 59(7): 103508.